



7683-031

SEQUENCE LISTING

RECEIVED

AUG 01 2001

TECH CENTER 1600/2900

(1) GENERAL INFORMATION:

- (i) APPLICANT: Ullrich, Axel  
Alves, Frauke
- (ii) TITLE OF INVENTION: MCK-10, A Novel Receptor Tyrosine Kinase
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Pennie & Edmonds
  - (B) STREET: 1155 Avenue of the Americas
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 10036-2711
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/153,397
  - (B) FILING DATE: 16-NOV-1993
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Coruzzi, Laura A.
  - (B) REGISTRATION NUMBER: 30,742
  - (C) REFERENCE/DOCKET NUMBER: 7683-031
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (212) 790-9090
  - (B) TELEFAX: (212) 869-9741/8864
  - (C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3962 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 321..3077

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGGGCCTGAG ACTGGGGTGA CTGGGACCTA AGAGAATCCT GAGCTGGAGG CCCCCGACAG  
 60

CTGCTCTCGG GAGCCGCCTC CCGACACCCG AGCCCCGCCG GCGCCTCCCG CTCCCGGCTC  
 120

CCGGCTCCTG GCTCCCTCCG CCTCCCCCGC CCCTCGCCCC GCCGCCGAAG AGGCCCCGCT  
 180

CCCGGGTCGG ACGCCTGGGT CTGCCGGGAA GAGCGATGAG AGGTGTCTGA AGGTGGCTAT  
 240

TCACTGAGCG ATGGGGTTGG ACTTGAAGGA ATGCCAAGAG ATGCTGCCCC CACCCCCTTA  
 300

GGCCCCGAGGG ATCAGGAGCT ATG GGA CCA GAG GCC CTG TCA TCT TTA CTG  
 350

Met Gly Pro Glu Ala Leu Ser Ser Leu Leu  
 1 5 10

CTG CTG CTC TTG GTG GCA AGT GGA GAT GCT GAC ATG AAG GGA CAT TTT  
 398

Leu Leu Leu Leu Val Ala Ser Gly Asp Ala Asp Met Lys Gly His Phe  
 15 20 25

GAT CCT GCC AAG TGC CGC TAT GCC CTG GGC ATG CAG GAC CGG ACC ATC  
 446

Asp Pro Ala Lys Cys Arg Tyr Ala Leu Gly Met Gln Asp Arg Thr Ile  
 30 35 40

CCA GAC AGT GAC ATC TCT GCT TCC AGC TCC TGG TCA GAT TCC ACT GCC

494  
 Pro Asp Ser Asp Ile Ser Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala  
           45                          50                          55  
 GCC CGC CAC AGC AGG TTG GAG AGC AGT GAC GGG GAT GGG GCC TGG TGC  
 542  
 Ala Arg His Ser Arg Leu Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys  
           60                          65                          70  
 CCC GCA GGG TCG GTG TTT CCC AAG GAG GAG GAG TAC TTG CAG GTG GAT  
 590  
 Pro Ala Gly Ser Val Phe Pro Lys Glu Glu Glu Tyr Leu Gln Val Asp  
       75                          80                          85                          90  
 CTA CAA CGA CTC CAC CTG GTG GCT CTG GTG GGC ACC CAG GGA CGG CAT  
 638  
 Leu Gln Arg Leu His Leu Val Ala Leu Val Gly Thr Gln Gly Arg His  
                           95                          100                          105  
 GCC GGG GGC CTG GGC AAG GAG TTC TCC CGG AGC TAC CGG CTG CGT TAC  
 686  
 Ala Gly Gly Leu Gly Lys Glu Phe Ser Arg Ser Tyr Arg Leu Arg Tyr  
                           110                          115                          120  
 TCC CGG GAT GGT CGC CGC TGG ATG GGC TGG AAG GAC CGC TGG GGT CAG  
 734  
 Ser Arg Asp Gly Arg Arg Trp Met Gly Trp Lys Asp Arg Trp Gly Gln  
           125                          130                          135  
 GAG GTG ATC TCA GGC AAT GAG GAC CCT GAG GGA GTG GTG CTG AAG GAC  
 782  
 Glu Val Ile Ser Gly Asn Glu Asp Pro Glu Gly Val Val Leu Lys Asp  
           140                          145                          150  
 CTT GGG CCC CCC ATG GTT GCC CGA CTG GTT CGC TTC TAC CCC CGG GCT  
 830  
 Leu Gly Pro Pro Met Val Ala Arg Leu Val Arg Phe Tyr Pro Arg Ala  
 155                          160                          165                          170  
 GAC CGG GTC ATG AGT GTC TGT CTG CGG GTA GAG CTC TAT GGC TGC CTC  
 878  
 Asp Arg Val Met Ser Val Cys Leu Arg Val Glu Leu Tyr Gly Cys Leu  
                           175                          180                          185  
 TGG AGG GAT GGA CTC CTG TCT TAC ACC GCC CCT GTG GGG CAG ACA ATG  
 926  
 Trp Arg Asp Gly Leu Leu Ser Tyr Thr Ala Pro Val Gly Gln Thr Met  
           190                          195                          200

TAT TTA TCT GAG GCC GTG TAC CTC AAC GAC TCC ACC TAT GAC GGA CAT  
 974  
 Tyr Leu Ser Glu Ala Val Tyr Leu Asn Asp Ser Thr Tyr Asp Gly His  
 205 210 215

ACC GTG GGC GGA CTG CAG TAT GGG GGT CTG GGC CAG CTG GCA GAT GGT  
 1022  
 Thr Val Gly Gly Leu Gln Tyr Gly Gly Leu Gly Gln Leu Ala Asp Gly  
 220 225 230

GTG GTG GGG CTG GAT GAC TTT AGG AAG AGT CAG GAG CTG CGG GTC TGG  
 1070  
 Val Val Gly Leu Asp Asp Phe Arg Lys Ser Gln Glu Leu Arg Val Trp  
 235 240 245 250

CCA GGC TAT GAC TAT GTG GGA TGG AGC AAC CAC AGC TTC TCC AGT GGC  
 1118  
 Pro Gly Tyr Asp Tyr Val Gly Trp Ser Asn His Ser Phe Ser Ser Gly  
 255 260 265

TAT GTG GAG ATG GAG TTT GAG TTT GAC CGG CTG AGG GCC TTC CAG GCT  
 1166  
 Tyr Val Glu Met Glu Phe Glu Phe Asp Arg Leu Arg Ala Phe Gln Ala  
 270 275 280

ATG CAG GTC CAC TGT AAC AAC ATG CAC ACG CTG GGA GCC CGT CTG CCT  
 1214  
 Met Gln Val His Cys Asn Asn Met His Thr Leu Gly Ala Arg Leu Pro  
 285 290 295

GGC GGG GTG GAA TGT CGC TTC CGG CGT GGC CCT GCC ATG GCC TGG GAG  
 1262  
 Gly Gly Val Glu Cys Arg Phe Arg Arg Gly Pro Ala Met Ala Trp Glu  
 300 305 310

GGG GAG CCC ATG CGC CAC AAC CTA GGG GGC AAC CTG GGG GAC CCC AGA  
 1310  
 Gly Glu Pro Met Arg His Asn Leu Gly Gly Asn Leu Gly Asp Pro Arg  
 315 320 325 330

GCC CGG GCT GTC TCA GTG CCC CTT GGC GGC CGT GTG GCT CGC TTT CTG  
 1358  
 Ala Arg Ala Val Ser Val Pro Leu Gly Gly Arg Val Ala Arg Phe Leu  
 335 340 345

CAG TGC CGC TTC CTC TTT GCG GGG CCC TGG TTA CTC TTC AGC GAA ATC  
 1406

Gln	Cys	Arg	Phe	Leu	Phe	Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	Glu	Ile	
			350					355					360			
TCC	TTC	ATC	TCT	GAT	GTG	GTG	AAC	AAT	TCC	TCT	CCG	GCA	CTG	GGA	GGC	
1454																
Ser	Phe	Ile	Ser	Asp	Val	Val	Asn	Asn	Ser	Ser	Pro	Ala	Leu	Gly	Gly	
		365					370					375				
ACC	TTC	CCG	CCA	GCC	CCC	TGG	TGG	CCG	CCT	GGC	CCA	CCT	CCC	ACC	AAC	
1502																
Thr	Phe	Pro	Pro	Ala	Pro	Trp	Trp	Pro	Pro	Gly	Pro	Pro	Pro	Thr	Asn	
	380					385					390					
TTC	AGC	AGC	TTG	GAG	CTG	GAG	CCC	AGA	GGC	CAG	CAG	CCC	GTG	GCC	AAG	
1550																
Phe	Ser	Ser	Leu	Glu	Leu	Glu	Pro	Arg	Gly	Gln	Gln	Pro	Val	Ala	Lys	
395					400					405					410	
GCC	GAG	GGG	AGC	CCG	ACC	GCC	ATC	CTC	ATC	GGC	TGC	CTG	GTG	GCC	ATC	
1598																
Ala	Glu	Gly	Ser	Pro	Thr	Ala	Ile	Leu	Ile	Gly	Cys	Leu	Val	Ala	Ile	
				415					420					425		
ATC	CTG	CTC	CTG	CTG	CTC	ATC	ATT	GCC	CTC	ATG	CTC	TGG	CGG	CTG	CAC	
1646																
Ile	Leu	Leu	Leu	Leu	Leu	Ile	Ile	Ala	Leu	Met	Leu	Trp	Arg	Leu	His	
			430					435					440			
TGG	CGC	AGG	CTC	CTC	AGC	AAG	GCT	GAA	CGG	AGG	GTG	TTG	GAA	GAG	GAG	
1694																
Trp	Arg	Arg	Leu	Leu	Ser	Lys	Ala	Glu	Arg	Arg	Val	Leu	Glu	Glu	Glu	
		445					450					455				
CTG	ACG	GTT	CAC	CTC	TCT	GTC	CCT	GGG	GAC	ACT	ATC	CTC	ATC	AAC	AAC	
1742																
Leu	Thr	Val	His	Leu	Ser	Val	Pro	Gly	Asp	Thr	Ile	Leu	Ile	Asn	Asn	
	460					465					470					
CGC	CCA	GGT	CCT	AGA	GAG	CCA	CCC	CCG	TAC	CAG	GAG	CCC	CGG	CCT	CGT	
1790																
Arg	Pro	Gly	Pro	Arg	Glu	Pro	Pro	Pro	Tyr	Gln	Glu	Pro	Arg	Pro	Arg	
475					480					485					490	
GGG	AAT	CCG	CCC	CAC	TCC	GCT	CCC	TGT	GTC	CCC	AAT	GGC	TCT	GCG	TTG	
1838																
Gly	Asn	Pro	Pro	His	Ser	Ala	Pro	Cys	Val	Pro	Asn	Gly	Ser	Ala	Leu	
				495					500					505		

CTG	CTC	TCC	AAT	CCA	GCC	TAC	CGC	CTC	CTT	CTG	GCC	ACT	TAC	GCC	CGT
1886															
Leu	Leu	Ser	Asn	Pro	Ala	Tyr	Arg	Leu	Leu	Ala	Thr	Tyr	Ala	Arg	
			510					515				520			
CCC	CCT	CGA	GGC	CCG	GGC	CCC	CCC	ACA	CCC	GCC	TGG	GCC	AAA	CCC	ACC
1934															
Pro	Pro	Arg	Gly	Pro	Gly	Pro	Pro	Thr	Pro	Ala	Trp	Ala	Lys	Pro	Thr
		525					530					535			
AAC	ACC	CAG	GCC	TAC	AGT	GGG	GAC	TAT	ATG	GAG	CCT	GAG	AAG	CCA	GGC
1982															
Asn	Thr	Gln	Ala	Tyr	Ser	Gly	Asp	Tyr	Met	Glu	Pro	Glu	Lys	Pro	Gly
	540					545					550				
GCC	CCG	CTT	CTG	CCC	CCA	CCT	CCC	CAG	AAC	AGC	GTC	CCC	CAT	TAT	GCC
2030															
Ala	Pro	Leu	Leu	Pro	Pro	Pro	Pro	Gln	Asn	Ser	Val	Pro	His	Tyr	Ala
555					560					565					570
GAG	GCT	GAC	ATT	GTT	ACC	CTG	CAG	GGC	GTC	ACC	GGG	GGC	AAC	ACC	TAT
2078															
Glu	Ala	Asp	Ile	Val	Thr	Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	Tyr
			575					580						585	
GCT	GTG	CCT	GCA	CTG	CCC	CCA	GGG	GCA	GTC	GGG	GAT	GGG	CCC	CCC	AGA
2126															
Ala	Val	Pro	Ala	Leu	Pro	Pro	Gly	Ala	Val	Gly	Asp	Gly	Pro	Pro	Arg
			590					595					600		
GTG	GAT	TTC	CCT	CGA	TCT	CGA	CTC	CGC	TTC	AAG	GAG	AAG	CTT	GGC	GAG
2174															
Val	Asp	Phe	Pro	Arg	Ser	Arg	Leu	Arg	Phe	Lys	Glu	Lys	Leu	Gly	Glu
		605					610					615			
GGC	CAG	TTT	GGG	GAG	GTG	CAC	CTG	TGT	GAG	GTC	GAC	AGC	CCT	CAA	GAT
2222															
Gly	Gln	Phe	Gly	Glu	Val	His	Leu	Cys	Glu	Val	Asp	Ser	Pro	Gln	Asp
	620					625					630				
CTG	GTC	AGT	CTT	GAT	TTC	CCC	CTT	AAT	GTG	CGT	AAG	GGA	CAC	CCT	TTG
2270															
Leu	Val	Ser	Leu	Asp	Phe	Pro	Leu	Asn	Val	Arg	Lys	Gly	His	Pro	Leu
635					640					645					650
CTG	GTA	GCT	GTC	AAG	ATC	TTA	CGG	CCA	GAT	GCC	ACC	AAG	AAT	GCC	AGC
2318															
Leu	Val	Ala	Val	Lys	Ile	Leu	Arg	Pro	Asp	Ala	Thr	Lys	Asn	Ala	Ser

655								660				665			
TTC	TCC	TTG	TTC	TCC	AGG	AAT	GAT	TTC	CTG	AAA	GAG	GTG	AAG	ATC	ATG
2366															
Phe	Ser	Leu	Phe	Ser	Arg	Asn	Asp	Phe	Leu	Lys	Glu	Val	Lys	Ile	Met
670								675				680			
TCG	AGG	CTC	AAG	GAC	CCC	AAC	ATC	ATT	CGG	CTG	CTG	GGC	GTG	TGT	GTG
2414															
Ser	Arg	Leu	Lys	Asp	Pro	Asn	Ile	Ile	Arg	Leu	Leu	Gly	Val	Cys	Val
685								690				695			
CAG	GAC	GAC	CCC	CTC	TGC	ATG	ATT	ACT	GAC	TAC	ATG	GAG	AAC	GGC	GAC
2462															
Gln	Asp	Asp	Pro	Leu	Cys	Met	Ile	Thr	Asp	Tyr	Met	Glu	Asn	Gly	Asp
700								705				710			
CTC	AAC	CAG	TTC	CTC	AGT	GCC	CAC	CAG	CTG	GAG	GAC	AAG	GCA	GCC	GAG
2510															
Leu	Asn	Gln	Phe	Leu	Ser	Ala	His	Gln	Leu	Glu	Asp	Lys	Ala	Ala	Glu
715								720				725			
GGG	GCC	CCT	GGG	GAC	GGG	CAG	GCT	GCG	CAG	GGG	CCC	ACC	ATC	AGC	TAC
2558															
Gly	Ala	Pro	Gly	Asp	Gly	Gln	Ala	Ala	Gln	Gly	Pro	Thr	Ile	Ser	Tyr
735								740				745			
CCA	ATG	CTG	CTG	CAT	GTG	GCA	GCC	CAG	ATC	GCC	TCC	GGC	ATG	CGC	TAT
2606															
Pro	Met	Leu	Leu	His	Val	Ala	Ala	Gln	Ile	Ala	Ser	Gly	Met	Arg	Tyr
750								755				760			
CTG	GCC	ACA	CTC	AAC	TTT	GTA	CAT	CGG	GAC	CTG	GCC	ACG	CGG	AAC	TGC
2654															
Leu	Ala	Thr	Leu	Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys
765								770				775			
CTA	GTT	GGG	GAA	AAT	TTC	ACC	ATC	AAA	ATC	GCA	GAC	TTT	GGC	ATG	AGC
2702															
Leu	Val	Gly	Glu	Asn	Phe	Thr	Ile	Lys	Ile	Ala	Asp	Phe	Gly	Met	Ser
780								785				790			
CGG	AAC	CTC	TAT	GCT	GGG	GAC	TAT	TAC	CGT	GTG	CAG	GGC	CGG	GCA	GTG
2750															
Arg	Asn	Leu	Tyr	Ala	Gly	Asp	Tyr	Tyr	Arg	Val	Gln	Gly	Arg	Ala	Val
795								800				805			
CTG	CCC	ATC	CGC	TGG	ATG	GCC	TGG	GAG	TGC	ATC	CTC	ATG	GGG	AAG	TTC

2798  
 Leu Pro Ile Arg Trp Met Ala Trp Glu Cys Ile Leu Met Gly Lys Phe  
                             815                            820                            825

ACG ACT GCG AGT GAC GTG TGG GCC TTT GGT GTG ACC CTG TGG GAG GTG  
 2846  
 Thr Thr Ala Ser Asp Val Trp Ala Phe Gly Val Thr Leu Trp Glu Val  
                             830                            835                            840

CTG ATG CTC TGT AGG GCC CAG CCC TTT GGG CAG CTC ACC GAC GAG CAG  
 2894  
 Leu Met Leu Cys Arg Ala Gln Pro Phe Gly Gln Leu Thr Asp Glu Gln  
                             845                            850                            855

GTC ATC GAG AAC GCG GGG GAG TTC TTC CGG GAC CAG GGC CGG CAG GTG  
 2942  
 Val Ile Glu Asn Ala Gly Glu Phe Phe Arg Asp Gln Gly Arg Gln Val  
                             860                            865                            870

TAC CTG TCC CGG CCG CCT GCC TGC CCG CAG GGC CTA TAT GAG CTG ATG  
 2990  
 Tyr Leu Ser Arg Pro Pro Ala Cys Pro Gln Gly Leu Tyr Glu Leu Met  
 875                            880                            885                            890

CTT CGG TGC TGG AGC CGG GAG TCT GAG CAG CGA CCA CCC TTT TCC CAG  
 3038  
 Leu Arg Cys Trp Ser Arg Glu Ser Glu Gln Arg Pro Pro Phe Ser Gln  
                             895                            900                            905

CTG CAT CGG TTC CTG GCA GAG GAT GCA CTC AAC ACG GTG TGAATCACAC  
 3087  
 Leu His Arg Phe Leu Ala Glu Asp Ala Leu Asn Thr Val  
                             910                            915

ATCCAGCTGC CCCTCCCTCA GGGAGTGATC CAGGGGAAGC CAGTGACACT AAAACAAGAG  
 3147

GACACAATGG CACCTCTGCC CTTCCCCTCC CGACAGCCCA TCACCTCTAA TAGAGGCAGT  
 3207

GAGACTGCAG GTGGGCTGGG CCCACCCAGG GAGCTGATGC CCCTTCTCCC CTTCTGGAC  
 3267

AACTCTCAT GTCCCCTTCC TGTTCTTCCT TCCTAGAAGC CCCTGTCGCC CACCCAGCTG  
 3327

GTCCTGTGGA TGGGATCCTC TCCACCCTCC TCTAGCCATC CCTTGGGGAA GGGTGGGGAG  
 3387



AAATATAGGA TAGACACTGG ACATGGCCCA TTGGAGCACC TGGGCCCCAC TGGACAACAC  
3447

TGATTCCTGG AGAGGTGGCT GCGCCCCAGC TTCTCTCTCC CTGTCACACA CTGGACCCCA  
3507

CTGGCTGAGA ATCTGGGGGT GAGGAGGACA AGAAGGAGAG GAAAATGTTT CCTTGTGCCT  
3567

GCTCCTGTAC TTGTCCTCAG CTTGGGCTTC TTCCTCCTCC ATCACCTGAA ACACTGGACC  
3627

TGGGGGTAGC CCCGCCCCAG CCCTCAGTCA CCCCCACTTC CCACTTGCAG TCTTGTAGCT  
3687

AGAACTTCTC TAAGCCTATA CGTTTCTGTG GAGTAAATAT TGGGATTGGG GGGAAAGAGG  
3747

GAGCAACGGC CCATAGCCTT GGGGTTGGAC ATCTCTAGTG TAGCTGCCAC ATTGATTTTT  
3807

CTATAATCAC TTGGGGTTTG TACATTTTTG GGGGGAGAGA CACAGATTTT TACACTAATA  
3867

TATGGACCTA GCTTGAGGCA ATTTTAATCC CCTGCACTAG GCAGGTAATA ATAAAGGTTG  
3927

AGTTTTCCAC AAAAAAAAAA AAAAAACCGG AATTC  
3962

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 919 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Pro Glu Ala Leu Ser Ser Leu Leu Leu Leu Leu Val Ala  
1 5 10 15

Ser Gly Asp Ala Asp Met Lys Gly His Phe Asp Pro Ala Lys Cys Arg  
20 25 30

Tyr	Ala	Leu	Gly	Met	Gln	Asp	Arg	Thr	Ile	Pro	Asp	Ser	Asp	Ile	Ser			
		35					40					45						
Ala	Ser	Ser	Ser	Trp	Ser	Asp	Ser	Thr	Ala	Ala	Arg	His	Ser	Arg	Leu			
	50					55					60							
Glu	Ser	Ser	Asp	Gly	Asp	Gly	Ala	Trp	Cys	Pro	Ala	Gly	Ser	Val	Phe			
	65				70					75					80			
Pro	Lys	Glu	Glu	Glu	Tyr	Leu	Gln	Val	Asp	Leu	Gln	Arg	Leu	His	Leu			
				85					90					95				
Val	Ala	Leu	Val	Gly	Thr	Gln	Gly	Arg	His	Ala	Gly	Gly	Leu	Gly	Lys			
			100					105					110					
Glu	Phe	Ser	Arg	Ser	Tyr	Arg	Leu	Arg	Tyr	Ser	Arg	Asp	Gly	Arg	Arg			
		115					120					125						
Trp	Met	Gly	Trp	Lys	Asp	Arg	Trp	Gly	Gln	Glu	Val	Ile	Ser	Gly	Asn			
	130					135					140							
Glu	Asp	Pro	Glu	Gly	Val	Val	Leu	Lys	Asp	Leu	Gly	Pro	Pro	Met	Val			
	145				150					155					160			
Ala	Arg	Leu	Val	Arg	Phe	Tyr	Pro	Arg	Ala	Asp	Arg	Val	Met	Ser	Val			
				165					170					175				
Cys	Leu	Arg	Val	Glu	Leu	Tyr	Gly	Cys	Leu	Trp	Arg	Asp	Gly	Leu	Leu			
			180					185					190					
Ser	Tyr	Thr	Ala	Pro	Val	Gly	Gln	Thr	Met	Tyr	Leu	Ser	Glu	Ala	Val			
		195					200					205						
Tyr	Leu	Asn	Asp	Ser	Thr	Tyr	Asp	Gly	His	Thr	Val	Gly	Gly	Leu	Gln			
	210					215					220							
Tyr	Gly	Gly	Leu	Gly	Gln	Leu	Ala	Asp	Gly	Val	Val	Gly	Leu	Asp	Asp			
	225				230					235					240			
Phe	Arg	Lys	Ser	Gln	Glu	Leu	Arg	Val	Trp	Pro	Gly	Tyr	Asp	Tyr	Val			
				245					250					255				
Gly	Trp	Ser	Asn	His	Ser	Phe	Ser	Ser	Gly	Tyr	Val	Glu	Met	Glu	Phe			
			260					265					270					
Glu	Phe	Asp	Arg	Leu	Arg	Ala	Phe	Gln	Ala	Met	Gln	Val	His	Cys	Asn			
		275					280					285						

Asn	Met	His	Thr	Leu	Gly	Ala	Arg	Leu	Pro	Gly	Gly	Val	Glu	Cys	Arg	
290						295					300					
Phe	Arg	Arg	Gly	Pro	Ala	Met	Ala	Trp	Glu	Gly	Glu	Pro	Met	Arg	His	
305					310					315					320	
Asn	Leu	Gly	Gly	Asn	Leu	Gly	Asp	Pro	Arg	Ala	Arg	Ala	Val	Ser	Val	
				325					330					335		
Pro	Leu	Gly	Gly	Arg	Val	Ala	Arg	Phe	Leu	Gln	Cys	Arg	Phe	Leu	Phe	
			340					345					350			
Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	Glu	Ile	Ser	Phe	Ile	Ser	Asp	Val	
		355					360					365				
Val	Asn	Asn	Ser	Ser	Pro	Ala	Leu	Gly	Gly	Thr	Phe	Pro	Pro	Ala	Pro	
	370					375					380					
Trp	Trp	Pro	Pro	Gly	Pro	Pro	Pro	Thr	Asn	Phe	Ser	Ser	Leu	Glu	Leu	
385					390					395					400	
Glu	Pro	Arg	Gly	Gln	Gln	Pro	Val	Ala	Lys	Ala	Glu	Gly	Ser	Pro	Thr	
				405					410					415		
Ala	Ile	Leu	Ile	Gly	Cys	Leu	Val	Ala	Ile	Ile	Leu	Leu	Leu	Leu	Leu	
			420					425					430			
Ile	Ile	Ala	Leu	Met	Leu	Trp	Arg	Leu	His	Trp	Arg	Arg	Leu	Leu	Ser	
		435					440					445				
Lys	Ala	Glu	Arg	Arg	Val	Leu	Glu	Glu	Glu	Leu	Thr	Val	His	Leu	Ser	
	450					455					460					
Val	Pro	Gly	Asp	Thr	Ile	Leu	Ile	Asn	Asn	Arg	Pro	Gly	Pro	Arg	Glu	
465					470					475					480	
Pro	Pro	Pro	Tyr	Gln	Glu	Pro	Arg	Pro	Arg	Gly	Asn	Pro	Pro	His	Ser	
				485					490					495		
Ala	Pro	Cys	Val	Pro	Asn	Gly	Ser	Ala	Leu	Leu	Leu	Ser	Asn	Pro	Ala	
			500					505					510			
Tyr	Arg	Leu	Leu	Leu	Ala	Thr	Tyr	Ala	Arg	Pro	Pro	Arg	Gly	Pro	Gly	
		515					520					525				
Pro	Pro	Thr	Pro	Ala	Trp	Ala	Lys	Pro	Thr	Asn	Thr	Gln	Ala	Tyr	Ser	
		530				535					540					

Gly	Asp	Tyr	Met	Glu	Pro	Glu	Lys	Pro	Gly	Ala	Pro	Leu	Leu	Pro	Pro	545	550	555	560
Pro	Pro	Gln	Asn	Ser	Val	Pro	His	Tyr	Ala	Glu	Ala	Asp	Ile	Val	Thr	565	570	575	
Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	Tyr	Ala	Val	Pro	Ala	Leu	Pro	580	585	590	
Pro	Gly	Ala	Val	Gly	Asp	Gly	Pro	Pro	Arg	Val	Asp	Phe	Pro	Arg	Ser	595	600	605	
Arg	Leu	Arg	Phe	Lys	Glu	Lys	Leu	Gly	Glu	Gly	Gln	Phe	Gly	Glu	Val	610	615	620	
His	Leu	Cys	Glu	Val	Asp	Ser	Pro	Gln	Asp	Leu	Val	Ser	Leu	Asp	Phe	625	630	635	640
Pro	Leu	Asn	Val	Arg	Lys	Gly	His	Pro	Leu	Leu	Val	Ala	Val	Lys	Ile	645	650	655	
Leu	Arg	Pro	Asp	Ala	Thr	Lys	Asn	Ala	Ser	Phe	Ser	Leu	Phe	Ser	Arg	660	665	670	
Asn	Asp	Phe	Leu	Lys	Glu	Val	Lys	Ile	Met	Ser	Arg	Leu	Lys	Asp	Pro	675	680	685	
Asn	Ile	Ile	Arg	Leu	Leu	Gly	Val	Cys	Val	Gln	Asp	Asp	Pro	Leu	Cys	690	695	700	
Met	Ile	Thr	Asp	Tyr	Met	Glu	Asn	Gly	Asp	Leu	Asn	Gln	Phe	Leu	Ser	705	710	715	720
Ala	His	Gln	Leu	Glu	Asp	Lys	Ala	Ala	Glu	Gly	Ala	Pro	Gly	Asp	Gly	725	730	735	
Gln	Ala	Ala	Gln	Gly	Pro	Thr	Ile	Ser	Tyr	Pro	Met	Leu	Leu	His	Val	740	745	750	
Ala	Ala	Gln	Ile	Ala	Ser	Gly	Met	Arg	Tyr	Leu	Ala	Thr	Leu	Asn	Phe	755	760	765	
Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys	Leu	Val	Gly	Glu	Asn	Phe	770	775	780	
Thr	Ile	Lys	Ile	Ala	Asp	Phe	Gly	Met	Ser	Arg	Asn	Leu	Tyr	Ala	Gly	785	790	795	800

Asp Tyr Tyr Arg Val Gln Gly Arg Ala Val Leu Pro Ile Arg Trp Met  
                             805                            810                            815  
 Ala Trp Glu Cys Ile Leu Met Gly Lys Phe Thr Thr Ala Ser Asp Val  
                             820                            825                            830  
 Trp Ala Phe Gly Val Thr Leu Trp Glu Val Leu Met Leu Cys Arg Ala  
                             835                            840                            845  
 Gln Pro Phe Gly Gln Leu Thr Asp Glu Gln Val Ile Glu Asn Ala Gly  
                             850                            855                            860  
 Glu Phe Phe Arg Asp Gln Gly Arg Gln Val Tyr Leu Ser Arg Pro Pro  
 865                            870                            875                            880  
 Ala Cys Pro Gln Gly Leu Tyr Glu Leu Met Leu Arg Cys Trp Ser Arg  
                             885                            890                            895  
 Glu Ser Glu Gln Arg Pro Pro Phe Ser Gln Leu His Arg Phe Leu Ala  
                             900                            905                            910  
 Glu Asp Ala Leu Asn Thr Val  
                             915

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3157 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 370..2934

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCACGAGCGG CACGAGTCCA TGATCTCTTT CCATCCTCCC TTTCCTGTTT GCTCACTTCT

60

TTTCTTGCTC ATCTTGGAGA CTGTGCAATC CCAGATTAAC TACAAACAGA GAAGAGCTGG  
120

TGATAGCTCC AGAGCTCAGA GAAAGGAGGT CTCTTTACAA GAAGTCTGGC TCTCAAAGCC  
180

TCCATCAAGG GAGACCTACA AGTTGCCTGG GGTTCAGTGC TCTAGAAAGT TCCAAGGTTT  
240

GTGGCTTGAA TTATTCTAAA GAAGCTGAAA TAATTGAAGA GAAGCAGAGG CCAGCTGTTT  
300

TTGAGGATCC TGCTCCACAG AGAATGCTCT GCACCCGTTG ATACTCCAGT TCCAACACCA  
360

TCTTCTGAG ATG ATC CTG ATT CCC AGA ATG CTC TTG GTG CTG TTC CTG  
408

Met Ile Leu Ile Pro Arg Met Leu Leu Val Leu Phe Leu  
1 5 10

CTG CTG CCT ATC TTG AGT TCT GCA AAA GCT CAG GTT AAT CCA GCT ATA  
456

Leu Leu Pro Ile Leu Ser Ser Ala Lys Ala Gln Val Asn Pro Ala Ile  
15 20 25

TGC CGC TAT CCT CTG GGC ATG TCA GGA GGC CAG ATT CCA GAT GAG GAC  
504

Cys Arg Tyr Pro Leu Gly Met Ser Gly Gly Gln Ile Pro Asp Glu Asp  
30 35 40 45

ATC ACA GCT TCC AGT CAG TGG TCA GAG TCC ACA GCT GCC AAA TAT GGA  
552

Ile Thr Ala Ser Ser Gln Trp Ser Glu Ser Thr Ala Ala Lys Tyr Gly  
50 55 60

AGG CTG GAC TCA GAA GAA GGG GAT GGA GCC TGG TGC CCT GAG ATT CCA  
600

Arg Leu Asp Ser Glu Glu Gly Asp Gly Ala Trp Cys Pro Glu Ile Pro  
65 70 75

GTG GAA CCT GAT GAC CTG AAG GAG TTT CTG CAG ATT GAC TTG CAC ACC  
648

Val Glu Pro Asp Asp Leu Lys Glu Phe Leu Gln Ile Asp Leu His Thr  
80 85 90

CTC CAT TTT ATC ACT CTG GTG GGG ACC CAG GGG CGC CAT GCA GGA GGT

696  
 Leu His Phe Ile Thr Leu Val Gly Thr Gln Gly Arg His Ala Gly Gly  
 95 100 105

CAT GGC ATC GAG TTT GCC CCC ATG TAC AAG ATC AAT TAC AGT CGG GAT  
 744

His Gly Ile Glu Phe Ala Pro Met Tyr Lys Ile Asn Tyr Ser Arg Asp  
 110 115 120 125

GGC ACT CGC TGG ATC TCT TGG CGG AAC CGT CAT GGG AAA CAG GTG CTG  
 792

Gly Thr Arg Trp Ile Ser Trp Arg Asn Arg His Gly Lys Gln Val Leu  
 130 135 140

GAT GGA AAT AGT AAC CCC TAT GAC ATT TTC CTA AAG GAC TTG GAG CCG  
 840

Asp Gly Asn Ser Asn Pro Tyr Asp Ile Phe Leu Lys Asp Leu Glu Pro  
 145 150 155

CCC ATT GTA GCC AGA TTT GTC CGG TTC ATT CCA GTC ACC GAC CAC TCC  
 888

Pro Ile Val Ala Arg Phe Val Arg Phe Ile Pro Val Thr Asp His Ser  
 160 165 170

ATG AAT GTG TGT ATG AGA GTG GAG CTT TAC GGC TGT GTC TGG CTA GAT  
 936

Met Asn Val Cys Met Arg Val Glu Leu Tyr Gly Cys Val Trp Leu Asp  
 175 180 185

GGC TTG GTG TCT TAC AAT GCT CCA GCT GGG CAG CAG TTT GTA CTC CCT  
 984

Gly Leu Val Ser Tyr Asn Ala Pro Ala Gly Gln Gln Phe Val Leu Pro  
 190 195 200 205

GGA GGT TCC ATC ATT TAT CTG AAT GAT TCT GTC TAT GAT GGA GCT GTT  
 1032

Gly Gly Ser Ile Ile Tyr Leu Asn Asp Ser Val Tyr Asp Gly Ala Val  
 210 215 220

GGA TAC AGC ATG ACA GAA GGG CTA GGC CAA TTG ACC GAT GGT GTG TCT  
 1080

Gly Tyr Ser Met Thr Glu Gly Leu Gly Gln Leu Thr Asp Gly Val Ser  
 225 230 235

GGC CTG GAC GAT TTC ACC CAG ACC CAT GAA TAC CAC GTG TGG CCC GGC  
 1128

Gly Leu Asp Asp Phe Thr Gln Thr His Glu Tyr His Val Trp Pro Gly  
 240 245 250

TAT GAC TAT GTG GGC TGG CGG AAC GAG AGT GCC ACC AAT GGC TAC ATT  
 1176  
 Tyr Asp Tyr Val Gly Trp Arg Asn Glu Ser Ala Thr Asn Gly Tyr Ile  
 255 260 265

GAG ATC ATG TTT GAA TTT GAC CGC ATC AGG AAT TTC ACT ACC ATG AAG  
 1224  
 Glu Ile Met Phe Glu Phe Asp Arg Ile Arg Asn Phe Thr Thr Met Lys  
 270 275 280 285

GTC CAC TGC AAC AAC ATG TTT GCT AAA GGT GTG AAG ATC TTT AAG GAG  
 1272  
 Val His Cys Asn Asn Met Phe Ala Lys Gly Val Lys Ile Phe Lys Glu  
 290 295 300

GTA CAG TGC TAC TTC CGC TCT GAA GCC AGT GAG TGG GAA CCT AAT GCC  
 1320  
 Val Gln Cys Tyr Phe Arg Ser Glu Ala Ser Glu Trp Glu Pro Asn Ala  
 305 310 315

ATT TCC TTC CCC CTT GTC CTG GAT GAC GTC AAC CCC AGT GCT CGG TTT  
 1368  
 Ile Ser Phe Pro Leu Val Leu Asp Asp Val Asn Pro Ser Ala Arg Phe  
 320 325 330

GTC ACG GTG CCT CTC CAC CAC CGA ATG GCC AGT GCC ATC AAG TGT CAA  
 1416  
 Val Thr Val Pro Leu His His Arg Met Ala Ser Ala Ile Lys Cys Gln  
 335 340 345

TAC CAT TTT GCA GAT ACC TGG ATG ATG TTC AGT GAG ATC ACC TTC CAA  
 1464  
 Tyr His Phe Ala Asp Thr Trp Met Met Phe Ser Glu Ile Thr Phe Gln  
 350 355 360 365

TCA GAT GCT GCA ATG TAC AAC AAC TCT GAA GCC CTG CCC ACC TCT CCT  
 1512  
 Ser Asp Ala Ala Met Tyr Asn Asn Ser Glu Ala Leu Pro Thr Ser Pro  
 370 375 380

ATG GCA CCC ACA ACC TAT GAT CCA ATG CTT AAA GTT GAT GAC AGC AAC  
 1560  
 Met Ala Pro Thr Thr Tyr Asp Pro Met Leu Lys Val Asp Asp Ser Asn  
 385 390 395

ACT CGG ATC CTG ATT GGC TGC TTG GTG GCC ATC ATC TTT ATC CTC CTG  
 1608



Thr	Arg	Ile	Leu	Ile	Gly	Cys	Leu	Val	Ala	Ile	Ile	Phe	Ile	Leu	Leu
		400					405					410			
GCC	ATC	ATT	GTC	ATC	ATC	CTC	TGG	AGG	CAG	TTC	TGG	CAG	AAA	ATG	CTG
1656															
Ala	Ile	Ile	Val	Ile	Ile	Leu	Trp	Arg	Gln	Phe	Trp	Gln	Lys	Met	Leu
415						420					425				
GAG	AAG	GCT	TCT	CGG	AGG	ATG	CTG	GAT	GAT	GAA	ATG	ACA	GTC	AGC	CTT
1704															
Glu	Lys	Ala	Ser	Arg	Arg	Met	Leu	Asp	Asp	Glu	Met	Thr	Val	Ser	Leu
430					435					440					445
TCC	CTG	CCA	AGT	GAT	TCT	AGC	ATG	TTC	AAC	AAT	AAC	CGC	TCC	TCA	TCA
1752															
Ser	Leu	Pro	Ser	Asp	Ser	Ser	Met	Phe	Asn	Asn	Asn	Arg	Ser	Ser	Ser
				450					455					460	
CCT	AGT	GAA	CAA	GGG	TCC	AAC	TCG	ACT	TAC	GAT	CGC	ATC	TTT	CCC	CTT
1800															
Pro	Ser	Glu	Gln	Gly	Ser	Asn	Ser	Thr	Tyr	Asp	Arg	Ile	Phe	Pro	Leu
			465					470					475		
CGC	CCT	GAC	TAC	CAG	GAG	CCA	TCC	AGG	CTG	ATA	CGA	AAA	CTC	CCA	GAA
1848															
Arg	Pro	Asp	Tyr	Gln	Glu	Pro	Ser	Arg	Leu	Ile	Arg	Lys	Leu	Pro	Glu
		480					485					490			
TTT	GCT	CCA	GGG	GAG	GAG	GAG	TCA	GGC	TGC	AGC	GGT	GTT	GTG	AAG	CCA
1896															
Phe	Ala	Pro	Gly	Glu	Glu	Glu	Ser	Gly	Cys	Ser	Gly	Val	Val	Lys	Pro
495						500					505				
GTC	CAG	CCC	AGT	GGC	CCT	GAG	GGG	GTG	CCC	CAC	TAT	GCA	GAG	GCT	GAC
1944															
Val	Gln	Pro	Ser	Gly	Pro	Glu	Gly	Val	Pro	His	Tyr	Ala	Glu	Ala	Asp
510					515					520					525
ATA	GTG	AAC	CTC	CAA	GGA	GTG	ACA	GGA	GGC	AAC	ACA	TAC	TCA	GTG	CCT
1992															
Ile	Val	Asn	Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	Tyr	Ser	Val	Pro
				530					535					540	
GCC	GTC	ACC	ATG	GAC	CTG	CTC	TCA	GGA	AAA	GAT	GTG	GCT	GTG	GAG	GAG
2040															
Ala	Val	Thr	Met	Asp	Leu	Leu	Ser	Gly	Lys	Asp	Val	Ala	Val	Glu	Glu
			545					550					555		

TTC	CCC	AGG	AAA	CTC	CTA	ACT	TTC	AAA	GAG	AAG	CTG	GGA	GAA	GGA	CAG
2088															
Phe	Pro	Arg	Lys	Leu	Leu	Thr	Phe	Lys	Glu	Lys	Leu	Gly	Glu	Gly	Gln
	560						565					570			
TTT	GGG	GAG	GTT	CAT	CTC	TGT	GAA	GTG	GAG	GGA	ATG	GAA	AAA	TTC	AAA
2136															
Phe	Gly	Glu	Val	His	Leu	Cys	Glu	Val	Glu	Gly	Met	Glu	Lys	Phe	Lys
575						580					585				
GAC	AAA	GAT	TTT	GCC	CTA	GAT	GTC	AGT	GCC	AAC	CAG	CCT	GTC	CTG	GTG
2184															
Asp	Lys	Asp	Phe	Ala	Leu	Asp	Val	Ser	Ala	Asn	Gln	Pro	Val	Leu	Val
590					595					600					605
GCT	GTG	AAA	ATG	CTC	CGA	GCA	GAT	GCC	AAC	AAG	AAT	GCC	AGG	AAT	GAT
2232															
Ala	Val	Lys	Met	Leu	Arg	Ala	Asp	Ala	Asn	Lys	Asn	Ala	Arg	Asn	Asp
				610					615					620	
TTT	CTT	AAG	GAG	ATA	AAG	ATC	ATG	TCT	CGG	CTC	AAG	GAC	CCA	AAC	ATC
2280															
Phe	Leu	Lys	Glu	Ile	Lys	Ile	Met	Ser	Arg	Leu	Lys	Asp	Pro	Asn	Ile
			625					630					635		
ATC	CAT	CTA	TTA	TCT	GTG	TGT	ATC	ACT	GAT	GAC	CCT	CTC	TGT	ATG	ATC
2328															
Ile	His	Leu	Leu	Ser	Val	Cys	Ile	Thr	Asp	Asp	Pro	Leu	Cys	Met	Ile
		640					645					650			
ACT	GAA	TAC	ATG	GAG	AAT	GGA	GAT	CTC	AAT	CAG	TTT	CTT	TCC	CGC	CAC
2376															
Thr	Glu	Tyr	Met	Glu	Asn	Gly	Asp	Leu	Asn	Gln	Phe	Leu	Ser	Arg	His
655						660					665				
GAG	CCC	CCT	AAT	TCT	TCC	TCC	AGC	GAT	GTA	CGC	ACT	GTC	AGT	TAC	ACC
2424															
Glu	Pro	Pro	Asn	Ser	Ser	Ser	Ser	Asp	Val	Arg	Thr	Val	Ser	Tyr	Thr
670					675					680					685
AAT	CTG	AAG	TTT	ATG	GCT	ACC	CAA	ATT	GCC	TCT	GGC	ATG	AAG	TAC	CTT
2472															
Asn	Leu	Lys	Phe	Met	Ala	Thr	Gln	Ile	Ala	Ser	Gly	Met	Lys	Tyr	Leu
				690					695					700	
TCC	TCT	CTT	AAT	TTT	GTT	CAC	CGA	GAT	CTG	GCC	ACA	CGA	AAC	TGT	TTA
2520															
Ser	Ser	Leu	Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys	Leu

705					710					715					
GTG	GGT	AAG	AAC	TAC	ACA	ATC	AAG	ATA	GCT	GAC	TTT	GGA	ATG	AGC	AGG
2568															
Val	Gly	Lys	Asn	Tyr	Thr	Ile	Lys	Ile	Ala	Asp	Phe	Gly	Met	Ser	Arg
		720					725					730			
AAC	CTG	TAC	AGT	GGT	GAC	TAT	TAC	CGG	ATC	CAG	GGC	CGG	GCA	GTG	CTC
2616															
Asn	Leu	Tyr	Ser	Gly	Asp	Tyr	Tyr	Arg	Ile	Gln	Gly	Arg	Ala	Val	Leu
	735					740					745				
CCT	ATC	CGC	TGG	ATG	TCT	TGG	GAG	AGT	ATC	TTG	CTG	GGC	AAG	TTC	ACT
2664															
Pro	Ile	Arg	Trp	Met	Ser	Trp	Glu	Ser	Ile	Leu	Leu	Gly	Lys	Phe	Thr
	750					755					760				765
ACA	GCA	AGT	GAT	GTG	TGG	GCC	TTT	GGG	GTT	ACT	TTG	TGG	GAG	ACT	TTC
2712															
Thr	Ala	Ser	Asp	Val	Trp	Ala	Phe	Gly	Val	Thr	Leu	Trp	Glu	Thr	Phe
				770					775						780
ACC	TTT	TGT	CAA	GAA	CAG	CCC	TAT	TCC	CAG	CTG	TCA	GAT	GAA	CAG	GTT
2760															
Thr	Phe	Cys	Gln	Glu	Gln	Pro	Tyr	Ser	Gln	Leu	Ser	Asp	Glu	Gln	Val
			785					790					795		
ATT	GAG	AAT	ACT	GGA	GAG	TTC	TTC	CGA	GAC	CAA	GGG	AGG	CAG	ACT	TAC
2808															
Ile	Glu	Asn	Thr	Gly	Glu	Phe	Phe	Arg	Asp	Gln	Gly	Arg	Gln	Thr	Tyr
		800					805					810			
CTC	CCT	CAA	CCA	GCC	ATT	TGT	CCT	GAC	TCT	GTG	TAT	AAG	CTG	ATG	CTC
2856															
Leu	Pro	Gln	Pro	Ala	Ile	Cys	Pro	Asp	Ser	Val	Tyr	Lys	Leu	Met	Leu
	815					820					825				
AGC	TGC	TGG	AGA	AGA	GAT	ACG	AAG	AAC	CGT	CCC	TCA	TTC	CAA	GAA	ATC
2904															
Ser	Cys	Trp	Arg	Arg	Asp	Thr	Lys	Asn	Arg	Pro	Ser	Phe	Gln	Glu	Ile
	830					835					840				845
CAC	CTT	CTG	CTC	CTT	CAA	CAA	GGC	GAC	GAG	TGATGCTGTC	AGTGCCTGGC				
2954															
His	Leu	Leu	Leu	Leu	Gln	Gln	Gly	Asp	Glu						
				850					855						
CATGTTTCCTA CGGCTCAGGT CCTCCCTACA AGACCTACCA CTCACCCATG CCTATGCCAC															

3014

TCCATCTGGA CATTTAATGA AACTGAGAGA CAGAGGCTTG TTTGCTTTGC CCTCTTTTCC  
3074

TGGTCACCCC CACTCCCTAC CCCTGACTCA TATATACTTT TTTTTTTTAC ATTAAAGAAC  
3134

TAAAAAAAAA AAAAAAAAAAG GCG  
3157

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 855 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Ile	Leu	Ile	Pro	Arg	Met	Leu	Leu	Val	Leu	Phe	Leu	Leu	Leu	Pro	1	5	10	15
Ile	Leu	Ser	Ser	Ala	Lys	Ala	Gln	Val	Asn	Pro	Ala	Ile	Cys	Arg	Tyr	20	25	30	
Pro	Leu	Gly	Met	Ser	Gly	Gly	Gln	Ile	Pro	Asp	Glu	Asp	Ile	Thr	Ala	35	40	45	
Ser	Ser	Gln	Trp	Ser	Glu	Ser	Thr	Ala	Ala	Lys	Tyr	Gly	Arg	Leu	Asp	50	55	60	
Ser	Glu	Glu	Gly	Asp	Gly	Ala	Trp	Cys	Pro	Glu	Ile	Pro	Val	Glu	Pro	65	70	75	80
Asp	Asp	Leu	Lys	Glu	Phe	Leu	Gln	Ile	Asp	Leu	His	Thr	Leu	His	Phe	85	90	95	
Ile	Thr	Leu	Val	Gly	Thr	Gln	Gly	Arg	His	Ala	Gly	Gly	His	Gly	Ile	100	105	110	
Glu	Phe	Ala	Pro	Met	Tyr	Lys	Ile	Asn	Tyr	Ser	Arg	Asp	Gly	Thr	Arg	115	120	125	
Trp	Ile	Ser	Trp	Arg	Asn	Arg	His	Gly	Lys	Gln	Val	Leu	Asp	Gly	Asn				

130					135					140					
Ser	Asn	Pro	Tyr	Asp	Ile	Phe	Leu	Lys	Asp	Leu	Glu	Pro	Pro	Ile	Val
145					150					155					160
Ala	Arg	Phe	Val	Arg	Phe	Ile	Pro	Val	Thr	Asp	His	Ser	Met	Asn	Val
				165					170					175	
Cys	Met	Arg	Val	Glu	Leu	Tyr	Gly	Cys	Val	Trp	Leu	Asp	Gly	Leu	Val
			180					185					190		
Ser	Tyr	Asn	Ala	Pro	Ala	Gly	Gln	Gln	Phe	Val	Leu	Pro	Gly	Gly	Ser
		195					200					205			
Ile	Ile	Tyr	Leu	Asn	Asp	Ser	Val	Tyr	Asp	Gly	Ala	Val	Gly	Tyr	Ser
	210					215					220				
Met	Thr	Glu	Gly	Leu	Gly	Gln	Leu	Thr	Asp	Gly	Val	Ser	Gly	Leu	Asp
225					230					235					240
Asp	Phe	Thr	Gln	Thr	His	Glu	Tyr	His	Val	Trp	Pro	Gly	Tyr	Asp	Tyr
				245					250					255	
Val	Gly	Trp	Arg	Asn	Glu	Ser	Ala	Thr	Asn	Gly	Tyr	Ile	Glu	Ile	Met
			260					265					270		
Phe	Glu	Phe	Asp	Arg	Ile	Arg	Asn	Phe	Thr	Thr	Met	Lys	Val	His	Cys
		275					280					285			
Asn	Asn	Met	Phe	Ala	Lys	Gly	Val	Lys	Ile	Phe	Lys	Glu	Val	Gln	Cys
	290					295					300				
Tyr	Phe	Arg	Ser	Glu	Ala	Ser	Glu	Trp	Glu	Pro	Asn	Ala	Ile	Ser	Phe
305					310					315					320
Pro	Leu	Val	Leu	Asp	Asp	Val	Asn	Pro	Ser	Ala	Arg	Phe	Val	Thr	Val
				325					330					335	
Pro	Leu	His	His	Arg	Met	Ala	Ser	Ala	Ile	Lys	Cys	Gln	Tyr	His	Phe
			340					345					350		
Ala	Asp	Thr	Trp	Met	Met	Phe	Ser	Glu	Ile	Thr	Phe	Gln	Ser	Asp	Ala
		355					360					365			
Ala	Met	Tyr	Asn	Asn	Ser	Glu	Ala	Leu	Pro	Thr	Ser	Pro	Met	Ala	Pro
	370					375					380				
Thr	Thr	Tyr	Asp	Pro	Met	Leu	Lys	Val	Asp	Asp	Ser	Asn	Thr	Arg	Ile

385					390						395				400
Leu	Ile	Gly	Cys	Leu	Val	Ala	Ile	Ile	Phe	Ile	Leu	Leu	Ala	Ile	Ile
				405					410					415	
Val	Ile	Ile	Leu	Trp	Arg	Gln	Phe	Trp	Gln	Lys	Met	Leu	Glu	Lys	Ala
			420					425					430		
Ser	Arg	Arg	Met	Leu	Asp	Asp	Glu	Met	Thr	Val	Ser	Leu	Ser	Leu	Pro
		435					440					445			
Ser	Asp	Ser	Ser	Met	Phe	Asn	Asn	Asn	Arg	Ser	Ser	Ser	Pro	Ser	Glu
	450					455					460				
Gln	Gly	Ser	Asn	Ser	Thr	Tyr	Asp	Arg	Ile	Phe	Pro	Leu	Arg	Pro	Asp
465					470					475					480
Tyr	Gln	Glu	Pro	Ser	Arg	Leu	Ile	Arg	Lys	Leu	Pro	Glu	Phe	Ala	Pro
				485					490					495	
Gly	Glu	Glu	Glu	Ser	Gly	Cys	Ser	Gly	Val	Val	Lys	Pro	Val	Gln	Pro
			500					505					510		
Ser	Gly	Pro	Glu	Gly	Val	Pro	His	Tyr	Ala	Glu	Ala	Asp	Ile	Val	Asn
		515					520					525			
Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	Tyr	Ser	Val	Pro	Ala	Val	Thr
	530					535					540				
Met	Asp	Leu	Leu	Ser	Gly	Lys	Asp	Val	Ala	Val	Glu	Glu	Phe	Pro	Arg
545					550					555					560
Lys	Leu	Leu	Thr	Phe	Lys	Glu	Lys	Leu	Gly	Glu	Gly	Gln	Phe	Gly	Glu
				565					570					575	
Val	His	Leu	Cys	Glu	Val	Glu	Gly	Met	Glu	Lys	Phe	Lys	Asp	Lys	Asp
			580					585					590		
Phe	Ala	Leu	Asp	Val	Ser	Ala	Asn	Gln	Pro	Val	Leu	Val	Ala	Val	Lys
		595					600					605			
Met	Leu	Arg	Ala	Asp	Ala	Asn	Lys	Asn	Ala	Arg	Asn	Asp	Phe	Leu	Lys
	610					615					620				
Glu	Ile	Lys	Ile	Met	Ser	Arg	Leu	Lys	Asp	Pro	Asn	Ile	Ile	His	Leu
625					630					635					640
Leu	Ser	Val	Cys	Ile	Thr	Asp	Asp	Pro	Leu	Cys	Met	Ile	Thr	Glu	Tyr

645								650				655			
Met	Glu	Asn	Gly	Asp	Leu	Asn	Gln	Phe	Leu	Ser	Arg	His	Glu	Pro	Pro
			660					665					670		
Asn	Ser	Ser	Ser	Ser	Asp	Val	Arg	Thr	Val	Ser	Tyr	Thr	Asn	Leu	Lys
		675					680					685			
Phe	Met	Ala	Thr	Gln	Ile	Ala	Ser	Gly	Met	Lys	Tyr	Leu	Ser	Ser	Leu
	690					695					700				
Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys	Leu	Val	Gly	Lys
705					710					715					720
Asn	Tyr	Thr	Ile	Lys	Ile	Ala	Asp	Phe	Gly	Met	Ser	Arg	Asn	Leu	Tyr
				725					730					735	
Ser	Gly	Asp	Tyr	Tyr	Arg	Ile	Gln	Gly	Arg	Ala	Val	Leu	Pro	Ile	Arg
			740					745					750		
Trp	Met	Ser	Trp	Glu	Ser	Ile	Leu	Leu	Gly	Lys	Phe	Thr	Thr	Ala	Ser
		755					760					765			
Asp	Val	Trp	Ala	Phe	Gly	Val	Thr	Leu	Trp	Glu	Thr	Phe	Thr	Phe	Cys
	770					775					780				
Gln	Glu	Gln	Pro	Tyr	Ser	Gln	Leu	Ser	Asp	Glu	Gln	Val	Ile	Glu	Asn
785					790					795					800
Thr	Gly	Glu	Phe	Phe	Arg	Asp	Gln	Gly	Arg	Gln	Thr	Tyr	Leu	Pro	Gln
				805					810					815	
Pro	Ala	Ile	Cys	Pro	Asp	Ser	Val	Tyr	Lys	Leu	Met	Leu	Ser	Cys	Trp
			820					825					830		
Arg	Arg	Asp	Thr	Lys	Asn	Arg	Pro	Ser	Phe	Gln	Glu	Ile	His	Leu	Leu
		835					840					845			
Leu	Leu	Gln	Gln	Gly	Asp	Glu									
	850					855									

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 3

(D) OTHER INFORMATION: /note= "Ala can be exchanged for any amino acid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asn Pro Ala Tyr

1

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Thr	Tyr	Ala	Xaa	Pro	Xaa	Xaa	Xaa	Pro	Gly
1				5					10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

His Arg Asp Leu Ala Ala



1

5

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGAATTCCCA YMGNRAYYTN RCNRCNMG  
28

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 6
- (D) OTHER INFORMATION: /note= "Xaa can be either Phe or Tyr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser Asp Val Trp Ser Xaa  
1 5

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGAATTCCYW YNSWGGTNTG SAGNST  
26

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

sp	His	Phe	Asp	Pro	Ala	Lys	Asp	Cys	Arg	Tyr	Ala	Leu	Gly	Met	Gln	A
	1			5				10				15				

Arg Thr Ile

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asn Thr Val

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 14 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

Pro Ala Met Ala Trp Glu Gly Glu Pro Met Arg His Asn Leu  
1 5 10

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 17 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

## Arg